



University of Groningen

A survey of complementaries in growth and location theories

Knaap, T.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1998

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Knaap, T. (1998). A survey of complementaries in growth and location theories. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

A survey of complementarities in growth and location theories

Thijs Knaap*
Groningen University

November 1998

Abstract

This survey looks at the use of the monopolistic competition framework in the theories of endogenous growth and economic geography.

* Faculty of Economics, P.O. Box 800, NL-9700 AV Groningen, The Netherlands. I thank Steven Brakman, Catrinus Jepma, Simon Kuipers, and participants of the 1997 NAKÉ day workshop for comments. Any remaining errors are mine. Correspondence preferably to t.knaap@eco.rug.nl

1. Introduction

In our ever-changing economy, few trends last so long that they may be used to characterize the developments from the industrial revolution until today. Yet over the centuries, two phenomena seem to have stood the test of time: every year, on average, economic output grows by a few percentage points (Romer 1986). And, through the years, economic activity has always agglomerated into small areas, instead of spreading out evenly (Krugman 1991a).

Not surprisingly, economic science has had some things to say about these two matters. However, in terms of models, the treatment has been rather upside-down. As for growth, the Solow (1956) model explains transitory adjustment processes, but the persistence of growth is an assumption rather than an outcome. Agglomerations are usually studied with models based on Von Thünen (1842), which show how the existence of a center affects the hinterland. The center, however, is also assumed rather than derived.

There exists an interesting connection between the deficiencies of these two approaches, and it is this connection that will be the theme of this survey. The inability of both models to generate the phenomena that seem so characteristic of real life is caused by the market form that is used. Both assume that economic activity is exclusively conducted by firms that are in full competition. This market form is in accordance with the firm's technical specification, namely, it is assumed that all firms are subject to constant returns to scale. It is easily seen that this assumption severely limits the possible outcomes.

If production is conducted under constant returns to scale, each separate factor in production faces decreasing returns. When growth is based on the accumulation of a subset of factors, this means that the economy cannot grow without bounds, and ends up in a steady state with zero growth.¹

In location theory, the spatial impossibility theorem of Starrett (1978)² states that a model with mobile agents on a closed, homogeneous space, facing a CRS³ production technology, can never explain the occurrence of agglomerations. Land rent will disperse economic activity without any countervailing force, because dividing up production over many locations leads to no loss in efficiency.⁴

1 For a detailed analysis, see Section 3.

2 The theorem is replicated in Fujita (1986).

3 When no ambiguity may arise, I use the expressions 'constant returns to scale,' 'constant returns,' and the acronym CRS interchangeably.

4 The clash between economic geography models and the need to specify the market structure is

Given these shortcomings of the CRS framework, it would seem tempting to use a wider class of firms, including those with increasing returns to scale. However, with the relinquishing of CRS, the assumption of full competition becomes untenable. The occurrence of other market forms greatly complicates the analysis, and allows for few analytical results.

Fortunately, a concise model of monopolistic competition introduced by Dixit and Stiglitz (1977) can be used to circumvent these problems. Among the upheaval that it caused in many areas of economics (see Buchanan and Yoon 1994), the model has allowed a new class of theories of growth and geography to be constructed. One of the characteristics of these new theories is that their main phenomena, growth and agglomeration, are now a consequence of the model, rather than an assumption.

This paper will briefly look at the monopolistic competition framework, and surveys the endogenous growth theory and economic geography in the light of it. It turns out that many interesting results in the two branches of literature can be attributed to the same fundamental properties of the monopolistic competition framework. The interplay between growth and geography is therefore not purely coincidental. While the models that show this were only recently made rigorous, their conclusions have been anticipated decades ago by such economists as Kaldor (1970) and Myrdal (1957):

“[...] the movements of labour, capital, goods and services do not by themselves counteract the natural tendency to regional inequality. By themselves, migration, capital movements and trade are rather the media through which the cumulative process evolves—upwards in the lucky regions and downwards in the unlucky ones.”(Myrdal 1957, p. 27)

The antiquated notion of ‘cumulative causation’ is revived today as a process caused by complementarities in the model.

I will first look into the nature of monopolistic competition and the complementarities that characterize it. This is done in Section 2. The findings are then used to provide a selective survey of endogenous growth theory (Section 3) and economic geography (Section 4). The two strands of literature are brought together in Section 5, and Section 6 concludes.

discussed at length in Krugman (1995).

2. Complementarities and the monopolistic competition framework

In the classical framework of economics, many important results are obtained under a broad set of assumptions. For instance, the propositions of welfare economics as they may be found in Arrow and Hahn (1971) or in Takayama (1985, p. 185), guarantee that in general, decentralized market outcomes are socially optimal.

The theory assumes, among others, that all producers of goods are in full competition. This assumption implies a number of important simplifications: under full competition, one producer's pricing decision does not influence the market in which she operates. Also, no producer makes a profit, and prices should equal marginal and average costs. These simplifications allow for simple pricing rules in the absence of strategic considerations. In this environment, a great number of analytical results may be derived.

As noted by Dixit and Stiglitz (1977, p. 297), the existence of a unique and optimal market equilibrium can be in jeopardy for at least three reasons, one of which is that potential economies of scale may not be used. Allowing economies of scale, however, means letting go of the CRS assumption. This alters the behavioral assumptions that are appropriate for the firms (Helpman 1984). Increasing returns imply, for instance, that the largest firm has the lowest average costs, and is able to push the smaller competitors off the market. Even if this may seem realistic for some sectors, it makes it much harder to derive analytical results.

There is a case for abandoning CRS however. The assumption tends to bend reality, and paints a world in which economic transactions are basically a zero-sum game. In a CRS economy, it is just as reasonable that all people divide their time over the same range of activities, as having each person specialize in one activity and allowing trade. Clearly, this outcome is unsatisfactory as a reflection of real economic activity. It goes as much against common sense as it goes against the founding words of economics as a science, dedicated to the productivity gains from dividing labor (Smith 1776, p.13).

The issue whether to assume CRS thus turned out to be rather crucial for a coherent model of general equilibrium, but unrealistic in practice. This left economists divided for a long time:

‘... there seem to be two traditions, which persist. On the one hand there are those who are so impressed by what has been done by the CRS method that they have come to live with it; on the other, those for whom scale economies are so important that they cannot bring themselves to leave them aside.’ (Hicks 1989, p. 12)

Among the efforts to bridge the gap was the work by Chamberlin (1933) and Robinson (1933), who sketched an alternative market form to the full competition implied by CRS. Their framework, monopolistic competition, held the promise of reconciling the two camps, but was rejected by most economists because of supposed inconsistency (Heijdra 1997). It was after the mathematical ramifications of the monopolistic competition framework had been studied by Spence (1976) and Dixit and Stiglitz (1977) that this alternative to CRS became widely used, especially in industrial economics, trade theory, growth theory and economic geography. Although it is ‘a very restrictive, indeed in some respects, a silly model’ (Krugman 1998, p. 164), it allows the economist to focus on the effects of increasing returns without worrying about strategic interactions between firms. The apparent arbitrariness of the model is not denied, but taken for granted, hoping that insights will extend beyond the model:

“Unfortunately, there are no general or even plausible tractable models of imperfect competition. The tractable models always involve some set of arbitrary assumptions about tastes, technology, behavior, or all three. This means that [...] one must have the courage to be silly, writing down models that are implausible in the details in order to arrive at convincing higher-level insights.” (Krugman 1995, pp. 14-15)

This section provides a short introduction to the Dixit-Stiglitz monopolistic competition framework. Before looking at the model itself, we will briefly discuss the problems that surround returns to scale in general, and the notion of externalities.

2.1 Returns to scale

A firm’s production possibilities are summarized in its production function. If for an amount A of a certain product a firm uses inputs, whose quantities are summarized in a vector \mathbf{B} , the correspondence between different values of A and \mathbf{B} defines the production function $f(\mathbf{B})$. For any \mathbf{B} , we can evaluate the returns to scale of the firm by looking at the point elasticity

$$\varepsilon_B = \frac{\partial f(\lambda \mathbf{B})}{\partial \lambda} \frac{1}{f(\mathbf{B})} \Big|_{\lambda=1}.$$

When ε_B is larger than one, there are increasing returns to scale. Note that ε_B is a function of the inputs \mathbf{B} . A firm can have increasing returns for all possible \mathbf{B} , but also for a limited set of values of \mathbf{B} .

On the level of the entire economy, increasing returns to scale are fairly undisputed. In this case, we can think of f as a nation’s production function, with \mathbf{B} indicating

the supply of labor and capital. Increasing returns have been attributed to the division of labor (Smith 1776) , splitting up complex production methods into multiple simple steps (Young 1928, Stigler 1951) , and the fact that technological knowledge, once produced, is nonexcludable (Romer 1990) . It would be a positive quality of any economic model to have the possibility of including increasing returns on the macro level.

However, much of today's macroeconomic theory is derived explicitly from micro-foundations (see, for instance, Romer 1993) . The occurrence of increasing returns at the micro-level spells trouble. Helpman (1984) shows that the modeler needs to specify a host of parameters to even start working: the conditions of firm entry, the heterogeneity of the good, and the type of market are just a few among them. The outcome of the model is highly dependent on these assumptions, for instance, do firms compete in a Bertrand- or a Cournot-market?

The simplest of these assumptions is that every sector is dominated by a single monopolist, who fully exploits the increasing returns. Apart from the question of realism, the presence of monopolists causes problems in a general-equilibrium model. One source of problems is the occurrence of monopoly rents: the model needs to specify how these rents are spent by the monopolist. In full competition, profits are zero by definition.

To avoid these issues altogether, one can assume that part of the returns to scale are external to the firm. The idea, originally from Marshall (1920) , separates internal economies ('those dependent on the resources of the individual houses of business engaged in it', p. 266) from external economies ('those dependent on the general development of the industry', p. 266). The distinction allows economies of scale to be incorporated in a consistent profit-maximizing framework, where firms perceive their situation as one of full competition. Between externalities, we can find two types (Scitovsky 1954): pecuniary externalities, those which are mediated by markets, and non-pecuniary externalities, those which are transmitted in an other way.

Non-pecuniary externalities use a production function, at the firm level, like $f(\mathbf{B}) = \tilde{f}(\mathbf{B}, X)$. Here, \mathbf{B} again are the inputs and X is industry output (Helpman 1984). Every single producer considers X as given, and controls only \mathbf{B} . But f may have increasing returns in \mathbf{B} and X together.

Pecuniary externalities are more subtle. It could be possible that a producer, by entering a market, increases the consumers' utility because of the increased variety that he/she provides. Although profit opportunities were the firm's original motive for entering, the variety effect may influence the perceived price level faced by the consumer, and alter the allocation of goods.

Using non-pecuniary externalities, it is possible to construct a general-equilibrium model that features increasing returns. Although this was indeed done (Chipman 1970), such models have not been used extensively. Especially the use of non-pecuniary externalities is thought suspect. By their nature, they are not observed so that one can assume anything about them. Any possible outcome can thus be ‘doctored’ into the model.

2.2 Monopolistic competition

The key difference between full competition and monopolistic competition⁵ is in the nature of the traded good. With full competition, the good is assumed to be homogeneous, and its price the only criterion of selection. With MC, consumers discern different varieties, and products from different producers are imperfect substitutes.⁶ Even if each individual producer faces increasing returns to scale in production, the largest producer is not always able to push smaller competitors out of the markets because substitution between products is limited.

In most applications of MC, consumer preferences are modelled as in Dixit and Stiglitz (1977)⁷. The quantities of goods x_i consumed are aggregated in a CES function,

$$U(x_1, \dots, x_n) = \left(\sum_{i=1}^n x_i^\theta \right)^{1/\theta} . \quad (1)$$

with $0 < \theta < 1$. By choosing suitable units of measurement for the different goods, we can abstain from adding scale parameters to the different x_i . It is clear that for each of the goods, an increase in the amount consumed will increase total utility. If we maximize (1) with respect to a budget constraint $\sum x_i p_i = E$, we find that

$$x_i = \frac{E}{q} \left(\frac{p_i}{q} \right)^{-\sigma}$$

where $\sigma = 1/(1 - \theta) > 1$, and the price index $q = \left(\sum p_j^{1-\sigma} \right)^{1/(1-\sigma)}$. So, each producer finds that she faces a demand elasticity σ .

5 We will use the acronym MC for ‘monopolistic competition’ from now on.

6 Chamberlin (1956, p. 56) suggests that such elements as ‘the conditions surrounding its sale’, trade marks and the seller’s reputation ‘may be regarded as [being purchased] along with the commodity itself.’

7 Weitzman (1994) shows that this model is much related to the Lancaster (1979) ‘spatial competition’ model, where each consumer has an ideal product and picks the one closest to it.

If every variety sells for the same price p , all are purchased in the same amount. In this case, formula (1) shows that utility is $n^{1/(\sigma-1)} E/p$. That is, an increase in variety brings an increase in utility even if the nominal budget remains the same. Helpman and Krugman (1985, p. 117) call this the ‘love-of-variety effect.’

The more varieties (n) there are, the less influence a single producer’s price exerts on the consumer’s real income. To completely eliminate every producer’s market power, it is often assumed that the range of goods $[0 \dots n]$ is continuous, and each producer is infinitely small. Though awkward, this assumption can be given some rigor. This is done in appendix A.

Producers are usually assumed to face a fixed cost F and a variable cost a per item produced. As the fixed cost per product declines with total production, they are subject to an increasing returns technology. However, because of the downward sloping consumer demand, output cannot grow indefinitely. Instead, producers maximize profits by setting marginal benefit equal to a . Facing demand generated by the utility function in (1) this pricing strategy results in a mark-up over marginal costs of size $1/\theta$. In equilibrium, all producers set the same price. The number of active producers adjusts so that discounted profits are just enough to recoup the initial investment F . With free entry, this means that n adjusts to drive profits to zero.

In an alternative interpretation of the same model, Ethier (1982) used the aggregator function in (1) as a production function. Output U is made with inputs x_i ; each input is produced by a single intermediate goods producer. The production function belongs to a class of firms that convert the intermediate goods into a final consumer good. These firms face constant returns to scale, as may be checked from (1), and are in full competition. The ‘love-of-variety effect’ from above has now become quite another thing: when entrance is free, there are increasing returns to scale at the economy’s macro level.

2.3 Complementarities

Matsuyama (1993, 1995) discusses complementarities, the notion that “two phenomena (or two actions, two activities) reinforce each other.” (1995, p. 702). Complementarities often arise in the MC framework.

As a specific example, assume that in an economy, people consume a single product that is made out of several intermediate goods with production function (1). That is, there are n different intermediate goods, and total production is U . This is the Ethier-setup from above. Assume also that intermediate-goods producers face fixed costs F and variable costs θx (the double use of parameter θ is for mathematical convenience). When there are L workers in the economy and there is free entry in the

intermediate sector, it can be computed that the number of producers in that sector will be

$$n^* = \frac{L}{\sigma F}$$

with σ defined as above. The per capita production is increasing in n^* , because of increasing returns to scale on the macro level. In fact, per capita production is $n^{1/(\sigma-1)}$.

Now if there exist two of these economies, with different intermediate goods, and they open up for trade, both economies will see the range of available intermediate goods increase. Because of this, both economies will experience an increase in production per capita. When the two economies interact, they are complementary to each other. This principle has been the basis for a large class of trade models, for instance in Helpman and Krugman (1985).

Hirschman (1958) discussed a related issue in the context of economic development. In his terminology, there exist *linkages* between different firms in a region. These linkages concern the input-output relations among the firms. Hirschman distinguishes *backward* linkages when a firm demands inputs from other firms, and *forward* linkages when a firm produces inputs for other firms. The conjecture is that with positive costs of transport for intermediate goods, linkages between firms can make an agglomeration stable.

In fact, the conjecture requires that linked firms are complementary to each other. It is true that in general, the arrival of a downstream firm can induce an upstream firm to expand. However, when this happens in a constant-returns world, the expansion has no effects on the original activities of the upstream firm, and the linkage is rather weak. But should the upstream firm exhibit increasing returns to scale, expansion means that it can now operate at a higher level of efficiency. In that case, the two firms are complementary.

2.4 Review, and a look ahead

To study a complex phenomenon, it can be necessary to make a number of assumptions that simplify the problem. We have argued that the CRS assumption fulfilled such a role in economics, as it allowed the derivation of a simple rule of conduct for firms, namely, marginal cost pricing. It also solved the problem of which market form would prevail, in favor of full competition.

We have also introduced an alternative framework, based on a different assumption: the MC setup. This setup is not any more general than full competition, the number of

assumptions has even increased. Yet it is an interesting alternative because it allows for complementarities and increasing returns to scale.

The short introduction above does not do justice to all the intricacies of MC, but that is not the point of this survey. Rather, we now want to look at the application of this framework to two fields, growth theory and economic geography. The application of MC to these fields has allowed a large number of innovations. Those in growth theory are discussed in the following Section, while those in economic geography are the subject of Section 4. The two strands of literature are brought together in Section 5.

3. Endogenous growth theory

In the introduction, we spoke briefly about the inability of traditional growth theory to explain lasting growth as an economic phenomenon. I will now substantiate these claims and introduce several alternatives that fall under the header of ‘new’ endogenous growth theory.

The MC framework introduced above does not play a pivotal role throughout endogenous growth theory. The new growth models were erected for a number of reasons, summarized by Romer (1994). Besides dissatisfaction with the inability of classical models to explain lasting growth, Romer identifies two other causes. One is the so-called convergence controversy: the (perceived) neoclassical prediction that poor countries must catch up with rich countries was disputed by data that became available around that time (Maddison 1982, Summers and Heston 1988). The other cause is the fact that the neoclassical model is at odds with a number of easily observable facts, facts which can only be explained if imperfect competition is incorporated.⁸ As the MC framework was the first to allow imperfect competition to be modelled in a concise way, it has been the framework of choice for a lot of endogenous growth models.

8 The facts are:

1. There are many firms in the economy, not one monopolist.
2. Discoveries are nonrival. This makes them different from other inputs.
3. Physical activities can be replicated; therefore production functions should be homogeneous of degree one.
4. Technological advance comes from things that people do. It does not occur by itself.
5. Many individuals have market power and earn monopoly rent on discoveries even though they are nonrival: information can be excludable.

Classical growth models are at odds with facts 4 and 5. Not all endogenous growth theories accommodate all these facts.

There are basically two ‘waves’ of models within the theory; the first wave (started by Romer 1986) describes growth as a process of ceaseless accumulation of factors. It is possible to retain the assumption of perfect competition in these models, using externalities. We will look at a sample model that employs the MC framework, though. The second wave (started by Romer 1990) explains growth by organized technological progress, and uses the MC framework together with an explicit sector for R&D.

We first briefly look at the exogenous (Solow-) growth model and compare it with some first-wave endogenous growth models. We then look at the second-wave models in Section 3.2.

3.1 Neoclassical and endogenous models of accumulation

3.1.1 The macro level

The neoclassical growth model was developed independently in Solow (1956) and Swan (1956), and the setup can be summarized quite concisely. The economy of a country uses two factors, L and K , and produces a single output. A proportion $(1 - s)$ is consumed⁹, the rest is used to increase K :

$$Y_t = F(K_t, L_t) \quad (2)$$

$$\dot{K}_t = sY_t \quad (3)$$

The aggregate production function F exhibits constant returns to scale, and the population of laborers L grows exponentially at rate n . To each factor L and K taken alone, the function has decreasing returns to scale. We may thus assume that the aggregate production function is a representation of an indeterminate number of firms that are in full competition.

The qualitative results of the model of course depend on the shape of F . Solow considers quite a number of different possibilities, but the one best remembered and usually quoted is when F has the Inada properties ($F_x \rightarrow \infty, 0$ as $x \rightarrow 0, \infty$ and $F(0, c) = F(c, 0) = 0$). Because of the CRS assumption, we may write this model in *per capita* terms by dividing both sides of (2) by L and substituting (3) in. This leads to the differential equation

$$\dot{k} = sf(k) - nk$$

9 The assumption of a fixed rate of saving can be relaxed without altering the basic results of the model. A model of intertemporal optimization was built by Cass (1965) and Koopmans (1965); the result may also be found in Barro and Sala-i-Martin (1995) and Rensman (1996).

where lowercase variables are per capita, and $f(k) = F(K/L, 1)$. By the Inada assumption, f exhibits decreasing returns to scale, so that the equation has a single solution k^* to which all time-paths must converge. This implies that there exists a level K/L at which at which the extra capital only just compensates the increase in population. This is the steady state to which the economy converges, and in which the growth in production per capita stops. The model is depicted in the left-hand panel

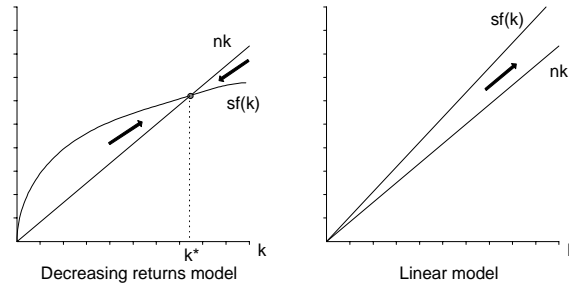


Figure 3.1: Direction of motion of k in two models of growth

of Figure 3.1. Capital per worker converges to the steady state level k^* from every initial level k_0 .

To stay in line with the empirical fact that the economy keeps growing, the neoclassical model is usually amended with exogenous technological growth. This growth is necessarily Harrod-neutral (for a proof, see Barro and Sala-i-Martin 1995, p. 54) and can be incorporated by substituting \hat{L}_t for L_t in (2), with $\hat{L}_t = A_t L_t$. Regular increases in A then result in a growing income per capita, even if the economy is in the steady state. If the rate of growth of A is assumed constant it is possible to estimate values for it for different countries using time series data. In another paper, I estimated exogenous growth for the U.S. to be 0.0180 [.0009] and for the Netherlands 0.0149 [.0021] (standard errors in brackets, Knaap 1997).

The neoclassical model highlights the process of capital accumulation in a closed economy and does not consider the interactions between several economies. It does make a prediction about the dispersion of capital per head over several closed economies, if these economies can all be described by the same production and investment functions: regardless of the initial level of capital, the economies will converge to the same equilibrium, and thus to the same level of K/L . This property of the model is known as the convergence property.

The temporary nature of growth in this model has to do with the fact that the factors that can be accumulated together face decreasing returns to scale. The more of these

accumulable factors are around, the less their added productivity is. This is an assumption of the model, and not necessarily a fact of life. The assumption was made because the neoclassical model also considers the factor labor, which cannot be accumulated by sheer economic means, and together the factors must exhibit CRS. For, if they do not exhibit CRS, the assumption of perfect competition is inappropriate.

On the premise that we will discuss the appropriate market structure in Section 3.1.2, let us now explore what would happen on a macro-level if *all* factors of production could be accumulated. This implies a return to the models proposed by Harrod (1939) and Domar (1946), who supposed that every addition to the stock of capital per worker allows production to be increased proportionally. Then the per capita stock of capital can never be too high, in the sense that additions to it are relatively unproductive. This can be seen when we substitute $F(K_t, L_t) = AK_t$ in formula (2) above. The accumulable resources in this case must be understood to include human capital and other production factors as well, besides capital in the narrow sense.

A graphical analysis of this linear model of production is in the right-hand panel of Figure 3.1. It is clear that if all factors can be accumulated, while the CRS condition still holds, we have specified a model of endogenous, ever-lasting growth.

An important point made by Rebelo (1991, p. 502) is that to achieve this result, not every part of the economy needs to have constant returns. It is sufficient that there exist a sector that uses a core of accumulable factors with a constant returns technology. This sector then becomes the economy's "engine of growth" as it pulls the rest of the economy.

We will illustrate this and other issues by considering the following two-sector model of an economy, taken from Barro and Sala-i-Martin (1995, p. 198):

$$C_t + \dot{K}_t + \delta_K K_t = A (v_t K_t)^{\alpha_1} (u_t H_t)^{\alpha_2} \quad (4)$$

$$\dot{H}_t + \delta_H H_t = B ((1 - v_t) K_t)^{\eta_1} ((1 - u_t) H_t)^{\eta_2} \quad (5)$$

A box-arrow sketch of this model is in Figure 3.2. The different colors of the arrows are used later; for now consider them all equal.

We see that there are two sectors, one with production function f (formula 5) and one with production function g (formula 4). Both sectors use two factors, K and H . In principle, both factors K and H can be accumulated. The variables v , u and C are control variables. The sectors differ in parameters α_i , η_i , A , and B and in the fact that the sector that produces K also produces consumption, C . Consumers solve the dynamic problem

$$\max_{u_t, v_t} \int_0^\infty \frac{C_t^{1-\sigma}}{1-\sigma} e^{-\rho t} dt$$

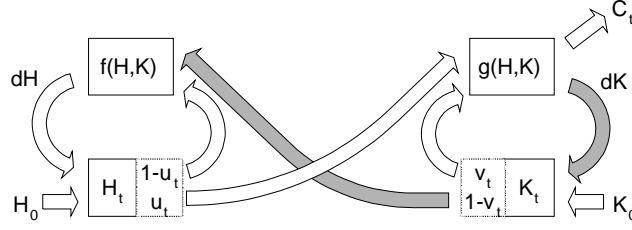


Figure 3.2: A box-arrow sketch of the two-sector model

given H_0 and K_0 and the parameters.

The complete model (4)-(5) is analyzed by Mulligan and Sala-i-Martin (1993). They derive the conditions under which this system can generate a steady state growth path, that is, a solution path where all variables grow at a constant rate. It turns out that this is only possible under the following condition:

$$(1 - \alpha_1)(1 - \eta_2) = \alpha_2 \eta_1 \quad (6)$$

A model whose parameters do not obey this condition either comes to rest at equilibrium levels of H and K or ‘explodes’, which means that it generates infinitely large state variables in finite time, and the objective integral becomes improper. This knife-edge condition on the parameters bothered Solow (1994) who discusses the value of α_1 in the AK model (see below). If that parameter is only slightly different than assumed, condition (6) is not satisfied and the endogenous growth results vanish. It causes him to call this type of theory “unpromising on theoretical grounds” (p. 51).

The model (4)-(5) has a number of well known special cases. We briefly list them below.

Example 1. *The AK model.* For this model, the sector on the left in Figure 3.2 is taken out. The other sector is assumed to have constant returns: $\alpha_2 = 0$, $\alpha_1 = v_t = 1$. Notice condition (6) is satisfied. This is a limiting case of the neoclassical Solow-Cass-Koopmans model with $f(K) = AK$, hence the name. The steady state solution is $\dot{C}/C = \dot{K}/K = (A - \delta_K - \rho)/\sigma$. The model does not have any transitional dynamics. The growth rate of C always remains positive under suitable parameters. \triangleleft

Example 2. *The engine of growth.* The two grey arrows in figure 3.2 are taken out. Both sectors have constant returns: $\eta_1 = 0$, $\eta_2 = 1$, $\alpha_1 = 1 - \alpha_2$, $\delta_K = \dot{K}_t = 0$. Here, K represents the invariant stock of non-reproducible, non-depreciating

capital goods (think of land, for instance) and H is the stock of factors that can be accumulated. Again, the model only has a steady state solution and lacks transitional dynamics. Rebelo (1991) shows that the solution is $\dot{C}/C = \alpha_1 \dot{H}/H$ which is equal to $\alpha_1 (B - \delta_H - \rho) / (1 - \alpha_1 (1 - \sigma))$. It is natural to designate the sector producing H as the engine of growth, as it is the constant returns accumulation of H that causes C to grow. \triangleleft

Example 3. *The Lucas model.* This is a slightly more general version of the ‘engine’ model from Example 2, analyzed in Lucas (1988). This time we take out only the middle grey arrow. The parameters are $\eta_1 = 0$, $\eta_2 = 1$, $\delta_{H,K} = 0$, $\alpha_1 + \alpha_2 > 1$. H is understood to be human capital and K is conventional capital. Thus capital goods play no role in the (constant returns) creation of human capital. The goods sector shows increasing returns. In fact, Lucas assumes constant returns plus an external effect of the average stock of human capital, so that a competitive equilibrium exists (more on this below). The optimal steady state growth rate of consumption (with zero population growth) is $\dot{C}/C = \dot{K}/K = (\frac{1-\alpha_1+\gamma}{1-\alpha_1} B - \rho)/\sigma$. Here, $\gamma = \alpha_1 + \alpha_2 - 1$, the size of the external effect. This shows that increasing returns are not essential for the resulting endogenous growth, as $\gamma = 0$ still permits a positive value for γ_C . \triangleleft

These models can be classified as to their stability. Because there are no transitional dynamics in the first two models, a small perturbation of the initial value has lasting effects. Because the growth rate of the accumulable factor is constant, the difference between the solutions starting in F and $F + \varepsilon$ grows exponentially (F is the initial value of the relevant state variable, K and H for the AK and the ‘engine’ model, respectively). A similar result holds for the Lucas model, although derivation of this result is not trivial. See Barro and Sala-i-Martin (1995, p. 184) and Mulligan and Sala-i-Martin (1993, p. 758).

3.1.2 The micro level

The models presented above pose a difficulty additional to the knife-edge condition on the parameters. If they include increasing returns to an accumulable factor, the usual fully competitive environment is no longer feasible; in other words, the set of supporting prices does not exist. We look at two approaches that have been used to circumvent this problem. One is to introduce increasing returns only at the level of the sector, and not of the firm. The sectorial returns take the shape of externalities. The other approach is to explicitly model the imperfect competition that arises because of the increasing returns.

Externalities We discussed externalities in Section 2.1 as a means to reconcile CRS and increasing returns. Some endogenous growth models use non-pecuniary externalities to do just this. We have already mentioned the use of externalities in the Lucas (1988) model, and we now look at the approach in Romer (1986). Because of a careful specification of the externality setup, the model does not suffer from the knife-edge condition (6).

The production function for a representative firm is $F(k_i, K, \mathbf{x}_i)$ with k_i the state of knowledge available to firm i and \mathbf{x}_i a vector of additional factors (capital, labor). The variable K is the aggregate level of knowledge $\sum_{i=1}^N k_i$ which can be used by all firms to some extent because knowledge is partly non-rival and non-excludable. It is assumed that F has constant returns to the factors k_i and \mathbf{x}_i , and increasing returns to all three factors. However, each firm takes the value of K as given when making its decisions. Output can be consumed or invested in k_i (\mathbf{x}_i is constant). The latter goes through the knowledge production function: $\dot{k}/k = g(I/k)$. The function g is increasing and bounded from above by a finite constant M . These conditions on g prevent the ‘explosion’ that the models above suffered from: a firm can never let its stock of knowledge grow at a faster rate than M so that k_i and K cannot reach infinity in finite time. Note that the g -functions above were usually linear in the state variable.

Romer finds that the socially optimal solution is different from the competitive solution because the latter does not take the external effects into account. Both solutions do generate endogenous growth, albeit that the rate of growth is larger in the optimal solution. The competitive solution is properly defined in all models that satisfy the above specification.

Monopolistic Competition As an alternative to the use of externalities above, Romer (1987) explicitly introduces markets that are monopolistically competitive; the model is very similar to that in Section 2.3. There exists an all-purpose capital good Z , which is transformed into a continuum of n^* intermediate goods; this is done by a continuum of firms (see appendix A). These intermediate goods are then used as inputs for the final good. The final good can again be added to Z or can be consumed. Consumers maximize utility (a function of consumption) intertemporally. The production function in the final goods sector is as described in section 2.2. An increasing number of intermediate inputs (n^*) increases output as in the example in Section 2.3. Varieties $x(i)$ are produced using an increasing returns production function.

The most important characteristic of this model is that output Y turns out to be a linear function of the stock Z . This is because the efficient scale of the intermediate producers does not change as Z changes, so n^* is linear in Z . As Y is linear in n^*

this means that the model behaves much as though it were the AK model above, and generates stable endogenous growth. It also suffers from the above-mentioned drawbacks, notably the fact that it is parameter-unstable. However, constant returns of Y to Z seem a little less “luck” (cf. Solow 1994, p. 51) than above, as they can be defended on economic grounds rather than just being mathematically convenient. Also, this model became the backbone of more advanced growth models. We will come across those models in the next section.

3.2 Growth through innovation

Above, economic growth was mostly brought about by an ever increasing supply of factors. In Romer’s (1987) model, an increase in the number of varieties played a role, but this increase was ‘free,’ *i.e.* no sacrifices needed to be made to discover the new varieties; the increase was a matter of efficient scale. Yet stylized fact #4 (footnote 8) specified that ‘technological advance comes from things that people do.’ The second wave of growth models thus concentrated on a situation where R&D absorbs resources and new varieties are discovered in return.

New varieties can be substitutes or complements to older ones. In growth theory parlance one thus distinguishes *horizontal* and *vertical* innovation. The term ‘horizontal innovation’ is from Grossman and Helpman (1991), and the first model in this direction was drafted by Judd (1985). It is replicated here.

It is assumed that consumers maximize an intertemporal CES utility function

$$U = \int_0^\infty e^{-\beta t} \left(\int_0^{V(t)} x(v, t)^\theta dv \right) dt. \quad (7)$$

The only factors of production are labor, which is constant at L , and the known range of varieties $V(t)$. For each variety there holds that one unit can be produced using one unit of labor. The range of varieties $V(t)$ grows through R&D, whose only input also is labor. It is assumed that $\dot{V} = L_{R\&D}/k$.¹⁰

There holds that $\theta < 1$, so that in equilibrium the quantities $x(v)$ are the same across varieties. Call this quantity y . The problem may then be written as

$$\begin{aligned} & \max_{0 \leq y \leq LV^{-1}} \int_0^\infty e^{-\beta t} y^\theta V dt \\ & \text{subject to } k\dot{V} = L - yV. \end{aligned}$$

¹⁰ Note that there is no uncertainty involved in research. This rather quaint assumption is maintained through much of the growth-through-innovation literature.

The solution (see Judd 1985) is that the economy converges to a stationary state where both y and V are constant. That is, there exists an optimal variety of goods, and once this variety has been attained innovation comes to a halt.

It is possible to see why innovation stops if we compare the problem to the basic monopolistic competition model of section 2.2. In that setup, an increase in that number of firms lowers each firm's profit margin. Profit is used to repay a fixed cost that is associated with entry. A situation of too many producers leads to profits that are too low to recoup the initial investment. Hence there exists an optimum number of producers. In this model, the fixed cost associated with entry is the labor that must be hired to conduct R&D. If that cost cannot be repaid because profit margins are too low, innovation stops.

Note that the MC market form is essential in this model because, as opposed to full competition, it allows producers to make a profit. Those profits can be used to pay off the initial R&D expenses. Without the possibility to price higher than marginal costs, innovation would never occur.

3.2.1 Horizontal innovation, endogenous growth

One way to keep the economy growing in the model above, is by lowering the costs of innovation as the number of varieties increases. If the outcome of the model should be a constant growth rate g of the number of varieties, and we know that $\dot{V} = L_{\text{R\&D}}/k$, then we can deduce

$$\begin{aligned} \frac{\dot{V}}{V} &= g = \frac{L_{\text{R\&D}}}{kV} \\ g \text{ constant} &\Rightarrow kV \text{ constant} \end{aligned}$$

So, if the R&D productivity parameter k^{-1} is proportional to the number of varieties V , we can have everlasting growth.

Romer (1990) presents an adapted version of his model in Romer (1991) that “emphasizes the importance of human capital in the research process” (p. S78). Like above, it features three sectors: R&D, intermediate and final goods. Knowledge has a rival component H and a non-rival component A ; the latter can be interpreted as the ‘state of technology’ and is allowed to grow without bounds.

In the production of final output Y , human capital H plays a role next to labor L and a continuum of intermediate goods $x(i)$:

$$Y(H_Y, L, x) = H_Y^\alpha \cdot L^\beta \cdot \int_0^A x(i)^{1-\alpha-\beta} di \quad (8)$$

(notice the similarity to formula 7 above). The stock of H is split up in a part H_Y that works in the final goods sector, and a part H_A that works in the R&D sector. The interval over which $x(i)$ is positive has size A , the level of technology. An increase in A , that is, a rise in the level of technology, does not render older types of the intermediate good obsolete. This is due to the additively separable form of (8).

In line with the derivation above, the technology used in the R&D sector is such that A changes according to

$$\dot{A} = \delta \cdot H_A \cdot A \quad (9)$$

This form is justified by claiming that a larger stock of knowledge will enhance current research possibilities. The model is closed by specifying that the stock of intermediate goods $K = \int_0^A x(i)di$ evolves according to $\dot{K}_t = Y_t - C_t$.

Romer's analysis shows that the model specified above yields unbounded endogenous growth. This is caused by the assumption of constant returns to scale in equation (9) above. With respect to this assumption, Romer writes:

“...in this sense, unbounded growth is more like an assumption than a result of the model. [...] Whether opportunities in research are actually petering out, or will eventually do so, is an empirical question that this kind of theory cannot resolve.”

3.2.2 Vertical innovation, endogenous growth

Aghion and Howitt (1992) consider a model of growth that features vertical innovation. Newer types of intermediates replace the older types, and therefore the model represents the concept of Creative Destruction introduced by Schumpeter (1942).

The economy consists of three sectors: the R&D sector, the intermediate goods sector and the sector that produces consumption goods. The trade-off in the economy is the decision how many workers are allotted to work in R&D instead of the intermediate goods sector. This number depends on the expected profitability of innovations.

A new intermediate good completely replaces the older type. The inventor is the only producer of the goods, and is thus allowed to earn some monopoly rents until the next innovation takes place. The time until the next innovation is random and exponentially distributed, and depends negatively on the number of people working in R&D. The marginal product of an extra R&D worker is decreasing, so that there exists an optimal number of people engaged in research and development.

There is only one kind of uncertainty in the model, namely the time of arrival of a new technology. The increase in the level of technology, caused by the invention, is

fixed. By defining a ‘period’ as the elapsed time between two innovations, the authors in effect make the monopoly rent earned off the inventions the random variable in the model.

Without being explicit about such things as the aggregate production function, Aghion and Howitt (1992) examine the motives for investing in R&D and find that, depending on the ‘arrival function’ of new technologies, there may exist a fluctuating or steady (possibly zero) number of researchers in the economy. Endogenous growth is implied as soon as there is a positive number of researchers active, and its rate is determined by both endogenous and exogenous variables.

3.3 Empirical tests

In this section we will mostly at empirical tests of the implications of the above models. As some of the results came out negatively, interest in the neoclassical Solow model was revived in the early 1990s. The results of such interest can be found in Mankiw, Romer and Weil (1992) and Nonneman and Vanhoudt (1996) .

As Pack (1994) notices, much early empirical research on endogenous growth models is conducted in the neoclassical framework. Thus, instead of testing the new growth theory directly, it is only used as a possible alternative when the Solow model fails. A first direct test of the theory is performed by Jones (1995b) , who tests the time-series predictions of new growth theory. The evidence is collected in two rounds.

The models of this section have the property that a permanent increase in investment causes a permanent increase in the economy’s rate of growth. Or, even stronger, the two variables are linearly related. This is easily seen with the *AK* and the Lucas model, as the rate of growth of capital and the rate of growth of consumption are the same. The result does not hold for the engine-of-growth model. Jones (1995b, p. 500) shows that the growth rates of selected OECD countries are stationary variables, whereas a unit root in the OECD investment rates can only be rejected in four out of 15 cases. Almost all investment rates show a positive trend. This contradicts the (supposed) linear relationship between investment rates and growth rates. Further time series estimations show that the effects of an increase in the investment rate can only be observed for eight years after the shock, much less than the proposed everlasting effect.

The testable proposition of the R&D-based models of section 3.2 is that the growth rate of an economy is linearly related to the number of people active in the R&D sector. Using data on the number of researchers in the U.S., Germany, Japan and France, Jones (1995b, p. 517) again shows a strong upward trend in these explanatory variables, whereas the rate of growth of their respective countries remains stationary.

These two results can be seen as a rejection of the testable propositions that came out of the endogenous growth models. Jones (1995a) proposes to ‘fix’ the R&D-based model by writing equation (9) in Romer’s (1990) model as

$$\dot{A} = \delta \cdot H_A^\lambda \cdot A^\phi$$

If $\lambda, \phi < 1$ then the model will no longer exhibit endogenous growth but instead settle down in an equilibrium. As Jones (1995a, p. 766) puts it, “... $\phi = 1$ represents a completely arbitrary degree of increasing returns and [...] is inconsistent with a broad range of time series data on R&D and TFP growth” (see also Romer’s quote in section 3.2.1). The model proposed by Jones (1995a) can best be seen as an extended version of the Solow (1956) setup, with all its asymptotic characteristics.

3.4 Review

We have seen that classical growth models that use the CRS paradigm explain growth through accumulation, but this growth cannot last forever without exogenous propelling. Accumulation-based models can explain lasting growth if they have constant returns to all accumulable factors. The micro-foundations for these models use externalities or an MC-setup.

The second wave of endogenous growth models explains growth not by accumulation of factors, but by technological progress. Virtually all these models use the MC framework.

Some critical notes can be placed about endogenous growth models. The scale-effects that they predict are not observed, and they are parameter-unstable. Despite the critical notes above, at the time of writing, endogenous growth theory is still very much alive. It turns out that the spirit of the models can be maintained while accommodating empirical facts (see Aghion and Howitt 1998, chapter 12). And the ability of the models to handle a number of questions that exogenous growth theory cannot answer (questions concerning the long run growth rate, for instance) has made them popular with empirical researchers.

4. Economic geography

Ironically, economic geography or location theory has been a rather peripheral field of study within economics. In part, the small amount of attention for issues of location can be attributed to the institutional, geographical and sociological factors that play such an important role in the problem. Yet over the years, many interesting results

have been obtained using methods of economics. We look at the foundations of location theory in Section 4.1. Then we turn to a new class of models that involve monopolistic competition and increasing returns in Section 4.2.

4.1 Foundations of location theory

The earliest theory of location can be divided in two branches (Greenhut 1956): least-cost theory, oriented on the supply side, and spatial competition theory, oriented on the demand side of the economy. The striking characteristic of least-cost theories is that they start by assuming a form of agglomeration; they do not explain why the agglomeration came about in the first place. This problem is tackled to some extent by spatial competition theories, as well as by the theories based on externalities and those that use increasing returns. We look at the different theories in chronological order.

Least-cost or land use theory starts with assuming that all demand in the economy is located at a single point. This can be a mining town demanding agricultural produce as in Von Thünen (1842), or a central business district in which all trade is conducted, as in Fujita (1986). Transportation is costly, and costs increase with distance from the center, r . From their production function and the costs of transport, suppliers can compute how much rent they want to pay as a function of r . This information is aggregated in a rent gradient, according to which the suppliers settle. The approach is refined by Weber (1909) to account for the location of raw materials, and Alonso (1964) adds, among other, endogenous lot size. Many models of urban structure still use this setup.

Spatial competition or locational interdependence theory, on the other hand, does not assume the existence of a center. Rather, (consumer) demand is distributed over locations and (zero-size) producers are looking for the optimal spot. With land rent out of the model, this approach clearly deals with questions of attraction and repulsion among different firms. The founding paper is Hotelling (1929), and shows that two producers of a homogeneous good will locate next to each other halfway a line with evenly spread consumers. This is not the socially optimal situation. Chamberlin (1956, pp. 260-265) shows that increasing the number of sellers in this problem will cause their dispersion, converging to the optimal dispersion as the number of sellers goes to infinity. Gabszewicz and Thisse (1986) provide a survey of this method.

The two approaches above may be combined. Lösch (1967) and Greenhut (1952) introduce profit-maximization as the relevant criterion. Given that demand and supply conditions may vary with location, this tends to make the problem less tractable. There exist fewer general rules on spatial dispersion than in the above, simplified,

analysis. An important limitation of both these approaches is the assumption that consumers do not change their location in response to the suppliers' whereabouts.

4.2 Models with endogenous agglomeration

The traditional location theory in the preceding section has said very little about the causes of agglomeration. Often, nonmarket externalities are thought to be an important factor in the creation of agglomerations. Such hard-to-measure concepts as informational and technical spillovers between firms, or in general informational exchanges between agents (Fujita and Thisse 1996, p.347) cause people to cluster together. The reason for clustering is the fact that the amount of spillovers between two firms is assumed to decline rapidly with distance. The spillovers are embodied in such acts as face-to-face talks and casual inspection of the other firm's production site. Nonmarket externalities are emphasized in Jacobs (1969) .

The problem with the above conjectures about the causes of agglomeration is that they are difficult to verify. Saying that agglomerations are caused by agglomeration economies is close to a tautology. The predictive power of the theory is therefore small. It is preferable to have a model where agglomerations are a result of more fundamental properties like the way people consume and produce.

It turns out that such a model can be constructed: the model includes pecuniary externalities as a cause for agglomeration and uses the monopolistic competition framework of section 2.2. The complementarities between different producers located at the same spot helps to make their combined presence an equilibrium. The principle, complementarity-induced agglomeration, was recognized by Krugman (1979) in a paper about monopolistic competition and international trade. In his model, where the MC setup was slightly different from above¹¹, when trade was prohibited but factors were mobile,

[. . .] there will be an incentive for workers to move to the region that already has the larger labor force. [. . .] In equilibrium all workers will have concentrated in one region or another. (p. 20)

It is not very difficult to see the agglomerative tendencies using the model from section 2.2. Suppose that there are two regions in which economies with an MC structure exist, and that trade is prohibited. The number of firms in each region is

11 Specifically, the subutility-function x^θ is replaced by a function $v(x)$. The elasticity of demand is now $-v'/v''x$, and it is assumed that the elasticity decreases in x (this does not happen with the x^θ form). The assumption leads to the result that wages are higher in the most populated region; with the x^θ form this is not true, unless the wages are corrected for the local price index.

linear in the number of inhabitants (Section 2.3); the aggregate price index faced by each inhabitant is (Section 2.2)

$$q = \left(\sum_{j=1}^n p_j^{1-\sigma} \right)^{1/(1-\sigma)} = (np^{1-\sigma})^{1/(1-\sigma)}$$

which is decreasing in n (this is due to the ‘love-of-variety-effect’). If inhabitants are given the choice where to live, they will move to the more populated region. Hence, agglomeration results naturally.

The first to design explicit models of location based on MC were Fujita (1988) and Rivera-Batiz (1988). These models featured agglomeration economies as well as land rents based on a least-cost framework (see Section 4.1 above). Despite the countervailing force of the rents, Rivera-Batiz shows that for some parameters, “[t]he economy’s population [...] ends up completely in city m ” (1988, p. 148).

Such complete agglomeration would become the hallmark property of a later class of models, then under the header of ‘new economic geography.’ The complementarities that cause agglomeration arise because of the MC framework, but they may travel through different markets. Ottaviano and Puga (1997) classify the models according to these media. They discern migration linkages, input-output linkages and intertemporal linkages.

A model that is based on migration linkages is presented by Krugman (1991a, 1991b). The Krugman model has two sectors, one mobile MC sector and a second sector which is immobile and fully competitive. The linkages work as follows: for a firm, it is preferable to be in a location with many inhabitants. This is because the inhabitants demand the firm’s product, and demand is increased if the firm is closer (because of transport costs). For people, on the other hand, it is preferable to be close to the largest concentration of firms, as we saw above, because they are subject to ‘love of variety.’ Being close to many firms lowers the price index they face and increases real wage.

Dependent on the relative size of the two sectors, transport costs, and substitution elasticities, the equilibrium may either be complete agglomeration of the mobile sector, or an even spread. A small change in the parameters may switch the equilibrium, so that ‘catastrophic’ changes are possible.

Venables (1996), in a model without labor, shows that it is possible that input-output linkages between firms fulfill the same role as a mobile workforce. Using a monopolistic competition setup for both an upstream and a downstream sector, Venables shows that it is possible that an increase in the size of one industry brings the other industry to a higher level of efficiency. The model’s conclusions remain the same in Krugman and Venables (1995), who extend the framework by collapsing the upstream

and downstream industries into one layer. The monopolistic competitive market structure is preserved by a specific form of the final demand function. Amiti (1997) shows that a similar outcome may be obtained without the use of an MC framework. In her model, a scale effect arises because of a pricing game that is played between firms in a sector. An increase in the number of firms has a negative effect on collusion and ups the sector's efficiency.

Aspects of factor accumulation can also serve as a medium for agglomerative tendencies. They are explored in Section 5 on dynamic economic geography.

Even though the mechanics, as well as the economic rationale of the above models are substantially different, there are some common characteristics that are worth spelling out. The most important outcome is that in all three models, the combination increasing returns - transport costs spells agglomeration. Both are a necessary factor. If there is no gain in size, then firms may as well split up and be spread out over space without any loss in efficiency. If transport costs are zero, then the whole concept of location does not matter in economic decisions (this is the spatial impossibility theorem referred to in Section 1).

The relation between transport costs and agglomeration tendencies is often found to be an 'inverted *U*' (Junius 1996, Ottaviano and Puga 1997, Venables 1996). At very high transport costs each region is self sufficient and no interaction takes place. At intermediate transport costs the above agglomeration effects are stronger, and at very low transport costs the 'centrifugal' forces congestion and factor market competition take over and firms spread out again.

4.3 Empirical tests

Due to its nonlinear structure, the empirical evaluation of the above economic geography models is quite laborious. Davis and Weinstein (1998a, 1998b) use the 'home market effect' discussed by Krugman (1980) to put a trade model to the test. Their model uses the Heckscher-Ohlin theory at the level of industries, and allows for a number of alternatives at the level of individual goods.

The home market effect may be observed if a region or a country has a large idiosyncratic demand for a particular good. In the Heckscher-Ohlin theory of trade, where the production structure is driven by factor endowments, such a region will in general be an importer of that good. The presence of an extraordinary level of demand does not affect the location of production. Even though local producers will satisfy the demand to some extent, they will not cover all of it, hence the importer status of the region.

In a world where producers only use one factor and compete within an MC framework,

things are different, and the home market effect surfaces: the region with the large demand for a specific good will be a net *exporter* of that good. The producers, rather than being driven by factor scarcity, realize that they are best off producing in only one location because of returns to scale, and prefer the region with the largest demand because of transport costs. Other regions are then serviced from this one, hence the net export result. In terms of the model, a large demand component is matched more than one for one by the region's production.

In the empirical specification, the authors lump together several goods into an 'industry,' explain industry location by endowments and look how the production of particular goods within the allotted industry production is distributed. In Davis and Weinstein (1998a), the data come from the OECD and concern national manufacturing production of the member countries. Results are meagre, in the sense that most trade can be explained by the traditional Heckscher-Ohlin model. When the model is estimated with data for regions in Japan (Davis and Weinstein 1998b), the results are quite different. When the model is estimated for each of the industries separately, two out of six feature a marked 'home market effect.'¹² The authors indicate that this could be the result of lower transport costs and greater factor mobility at the regional level.

The tests by Davis and Weinstein must be seen as a first coarse investigation into the relevance of economic geography models. Many assumptions are made: at what level of aggregation do 'industries' stop, and do we observe 'goods' where the market is monopolistically competitive? What goods form an industry together? How do arbitrary definitions of regions and groups of goods affect the results? It seems wise to defer judgement until more empirical evidence is available.

5. Dynamic economic geography

This review paper has shown how the monopolistic competition model (Section 2) made possible new ways of modelling economic growth (Section 3) and economic geography (Section 4). So far however, we have discussed growth models without an explicit geographical dimension, and static geography models. It is only natural to combine the two strands of the literature, which are based on the same framework.

Some research has indeed been done in this direction. We survey it in this section, organizing the models along the type of medium that is used to transmit the agglomeration linkages (these media were discussed in Section 4.2). We start with a model by

¹² These industries include: transportation equipment, general machinery, electrical machinery and precision instruments. Indeed these skill-intensive goods seem to be among the ones where differentiation is possible, as opposed to manufactured bulk goods.

Martin and Ottaviano (1996b), where the linkage runs through the R&D sector in a way reminiscent of Krugman and Venables (1995), in Section 5.1. A model by Baldwin and Forslid (1997) where the labor market is the medium, is the subject of Section 5.2. Other models are in 5.3.

5.1 Agglomeration through the R&D sector

The model by Martin and Ottaviano (1996b) has two locations and three sectors: a full-competition agricultural sector, an MC industrial sector and a sector for R&D. It is the latter sector that is most interesting.

The R&D sector is fully competitive. The output of the sector is patents; each patent can be used to manufacture a variety in the industrial sector, the total number of varieties is n . The productivity of the R&D sector increases as n gets larger. These qualities are similar to the Romer-Grossman-Helpman models of Section 3.2.1. The only input to the R&D sector is the composite good D that is the output of the industrial sector. This creates a linkage between the two sectors akin to the linkages in Krugman and Venables (1995) and Venables (1996). Wherever firms from the industrial sector are abundant, the costs of R&D are low. And wherever R&D is conducted, the demand for industrial goods is higher. The linkage causes agglomeration of industrial and R&D firms in the same location.

Consumers in this model maximize an intertemporal utility function that depends on the consumption of the agricultural and the industrial good. The model has two types of solutions. In one solution, both locations have exactly the same number of industrial producers. R&D is conducted in both locations. This solution is unstable. The other solution has all R&D taking place in one location, where also the majority of the industrial producers are active.

It is in the second, unbalanced, solution the rate of growth is higher. This is intuitive: if industrial producers are spread evenly the industrial composite costs the same in both locations, say, c . In case of an imbalance, there always is a location in which the composite is cheaper than c . Because R&D uses only the composite, an even spread of the industrial producers maximizes production costs and minimizes growth.

An important conclusion of the model by Martin and Ottaviano (1996b) is that the rate of growth influences the location decision, and the location decision influences the rate of growth. This puts models in which both are treated separately at a disadvantage. The fact that the interaction causes agglomeration of industrial activity is in line with the quote by Myrdal on page 3.

5.2 Agglomeration through the labor market

The model in the previous section assumed that there is no migration between the locations. At the other extreme, migration is the cornerstone in the model by Baldwin and Forslid (1997) , just as it is in Krugman (1991a, 1991b) .

The assumptions are roughly the same as above, except that the R&D sector now uses only labor as an input. Again, the input requirement decreases as the stock of knowledge gets larger. However, the spillovers are only regional; the stock of knowledge consists of the number of firms in ones own location only.

In the long run, the linkage now works as follows: wherever the most firms are is where the consumer price index is lowest. Personnel has an incentive to move to this location. So do all firms in the (competitive) R&D sector, because the costs of R&D depend negatively on the available pool of knowledge (in this case, the number of firms). On the other hand, where most people are is where firms like to be because of the demand that people exercise, and because of the larger labor market that the firms can draw from.

Again, there are two types of equilibrium in this model. In one, all activity is evenly divided between locations, and both locations grow at the same speed. The other equilibrium has all R&D and most labor and industrial firms in one location, the other deprived of most activity.

It turns out that the first equilibrium (the even spread) is very unstable, even at prohibitive trade costs. This was not the case in the Krugman (1991a, 1991b) models. Contrary to the static economy, the dynamic economy will agglomerate into one location for all possible parameters.

In this model, the R&D sector does not constitute a part of the linkages, as it did above. However, it does react to the outcome. In the long term, all R&D is concentrated in the agglomeration, because it is the cheaper place to work. This does not necessarily affect the location of the industrial firms developed by the R&D sector, as the patents are valid in both locations. Thus, the R&D sector reacts to the linkages, but is not a part of it.

The interplay between growth and location shows up in this model as well. When all R&D is done in the same location, all R&D firms add to the same stock of knowledge. This leads to faster rates of growth than if the advances are divided over two separate stocks of knowledge, because the efficiency of the R&D sector increases with K .

5.3 Other models of growth and geography

The models in this section integrate growth and location theories as above, but they differ from the models in Sections 5.1 and 5.2 in the fact that growth and location do not interact. As we saw above, the interaction of growth and location was the most interesting aspect of the hybrid models, so we do not analyze the models in this Section to a great extent.

Martin and Ottaviano (1996a) develop a model where migration does not occur. There are three sectors, agriculture, industrial and R&D. The MC industrial sector uses patents as in Section 3.2.1. The competitive R&D sector that develops the patents uses labor and the pool of knowledge. Patents can be used in any location and are not subject to transport costs. If the R&D sector has access to all knowledge in the economy (global spillovers), then R&D is conducted in both locations. If there are only local spillovers, the R&D sector agglomerates. The developed firms will be set up in both locations, though.

The model is a first attempt to merge theories of growth and location. The structure of the economy (industrial and agricultural production at the two locations) is so rigid that it does not change much under the different growth regimes, so that the interaction is limited to the location of the R&D sector.

Englmann and Walz (1995) construct a model with two locations without transport costs. The geographic structure plays a role however, because the knowledge pool is different between the two regions. This leads to a situation with nontraded inputs, where each location has its own intermediates. The initially larger region becomes the industrial center, whereas the other becomes a peripheral region. If there are interregional knowledge spillovers, so that inputs still are not traded but R&D *can* use them, many solutions become possible.

In this model, devoid of transport costs, it is the size of the knowledge pools that steers the regional development. Knowledge pools contain nontraded inputs, so that the factor that causes agglomeration is not traded itself. Though interesting, this is fundamentally different from the models of Section 4 and is the subject of another branch of literature (see, for instance Rivera-Batiz and Romer 1991) .

6. Conclusions

In this survey paper, we introduced the monopolistic competition framework as the foundation of two new strands of literature, on the one hand endogenous growth theory, and on the other hand economic geography. Both theories use the fact that

MC allows scale economies to be used in a model of general equilibrium.

In our survey of endogenous growth, we showed that early models were based on the endless accumulation of resources, as in exogenous growth models. Later versions stressed technological progress as the source of growth. Progress can take the form of horizontal innovations and vertical innovations.

In the literature on economic geography, linkages between firms and consumers, and between firms themselves, play an important role. The different models can be classified as to the type of linkage they use. Most models predict a dramatic agglomeration at certain parameter values.

Because both strands of literature rest on the same foundation, and describe related phenomena, it is only logical to incorporate the two. We surveyed several attempts to that end. It turns out that the interplay between growth and location upsets the predictions of either literature by itself. Stable equilibria in static geography models turn out to be unstable in a dynamic context; the rate of growth again is influenced by the location pattern, which depends on initial values.

Studies that investigate the empirical value of both literatures are not overly enthusiastic. Whereas CRS-based theory stands up to the data in a reasonable way, many effects predicted by MC are not measured at all. However, this may be due to a lack of testing methodology capable of dealing with the nonlinear nature of the models. Tests can only be conducted on specific linear predictions of the model. It is unclear to which extent a refutation of such a prediction constitutes a problem for the whole body of theory.

It seems that the combination of endogenous growth theory and economic geography is a promising field of research. The scattered results available so far indicate that more work needs to be done before any swaying conclusions can be drawn.

]

Appendix A:

A continuum of goods

The derivation of the equilibrium in the monopolistic competition framework holds in general ‘when n is large.’ This can be an awkward assumption; do we really need, in economic terms, an endless array of goods to work with this model?

The usual interpretation is that really, all we need is to be able to refine and differentiate goods enough. The range can remain the same, but we ought to be able to divide goods into as many different subtypes as we need. Mathematically, this means that we look at a continuum of goods $x(j)$ defined on a real interval $[0, n]$. In principle, each good $x(j)$ with $j \in [0, n]$ can be identified as a different variety. Quantities of goods, however, are only defined over intervals of j . The quantity $x(3) = 1$ is meaningless, but $x(j) = 1$ for all $j \in [0.1, 0.2]$ is a positive quantity.

How do our maximand U and the budget restriction change when we work with a continuum of goods? They can be derived as limiting cases of their discrete versions.

Suppose we call all the goods $x(j)$ with $0 \leq j < n_1$ good 1, all the goods with $n_1 \leq j < n_2$ good 2, and introduce a set of numbers $S = \{n_0, n_1, n_2, \dots, n_Q\}$ like this, with $n_0 = 0$ and $n_Q = n$. If two goods belong to the same interval, they are purchased in the same amount and priced the same.¹³ With this set, we are back in the discrete goods setup. There holds

$$U = \left[\sum_{i=1}^Q (n_i - n_{i-1}) x(i)^\theta \right]^{\frac{1}{\theta}}$$

$$E \geq \sum_{i=1}^Q (n_i - n_{i-1}) x_i p_i.$$

For any properly defined set S , these formulae can be rewritten as

$$U = \left[\int_0^n x(i)^\theta di \right]^{\frac{1}{\theta}} \quad (10)$$

$$E \geq \int_0^n x(i) p(i) di \quad (11)$$

We see that there are two ways in which the number of goods can increase. By picking a larger set S , we refine the definition of the goods, and allow for more price and quantity differentiation. By increasing n , the range of goods is increased with the introduction of new varieties that can be purchased *instead of* the older set.

The monopolistic competition setup is usually introduced as in formulas (10) and (11), without a specific set S defined. To retrieve the results that hold in the integer case, however, we need to imagine such a set ourselves.

Suppose we want to maximize function U from formula (10) under the restriction

¹³ That is, we have $x(i)$ and $x(j)$ with $n_{k-1} \leq i < n_k$ and $n_{k-1} \leq j < n_k$, and both are purchased in the amount x_k .

(11). The problem can be written as a Lagrangian,¹⁴

$$\max_{\{x(i)|i \in [0,n]\}} \left[\int_0^n x(i)^\theta di \right] - \lambda \left[\int_0^n x(i) p(i) di - E \right]$$

The problem is hard to solve when we stick with the integral notation, but we can imagine that the differentiation between goods only goes as far as a set \mathcal{S} , which we do not specify. We may then write the maximand as

$$\mathcal{L} = \left[\sum_{i=1}^Q (n_i - n_{i-1}) x(i)^\theta \right] - \lambda \left[\sum_{i=1}^Q (n_i - n_{i-1}) x(i) p(i) - E \right].$$

Differentiate with respect to $x(i)$ and set equal to zero to find

$$\begin{aligned} \beta (n_i - n_{i-1}) x(i)^{\theta-1} - \lambda (n_i - n_{i-1}) p(i) &= 0 \Rightarrow \\ \theta x(i)^{\theta-1} &= \lambda p(i). \end{aligned}$$

Note that we may divide by $(n_i - n_{i-1})$ because the requirements for \mathcal{S} have it greater than zero. Because λk does not vary with i , we may write that for all i ,

$$x(i) p(i)^{\frac{1}{\theta-1}} = \text{constant}.$$

If we substitute this into formula (11) we get that

$$x(i) = \frac{E p(i)^{-\sigma}}{\int_0^n p(j)^{1-\sigma} dj}$$

where $\sigma = 1/(1 - \theta) > 1$.

References

- Aghion, P. and Howitt, P.: 1992, A model of growth through creative destruction, *Econometrica* **60**(2), 323–51.
- Aghion, P. and Howitt, P.: 1998, *Endogenous Growth Theory*, The MIT Press, Cambridge, MA.
- Alonso, W.: 1964, *Location and Land Use: Toward a General Theory of Land Rent*, Harvard University Press, Cambridge, MA.
- Amiti, M.: 1997, Regional specialisation and technological leapfrogging. mimeo, La Trobe University, Victoria, Australia.

¹⁴ We momentarily omit the exponent $1/\theta$, which does not change the outcome of the maximization.

- Arrow, K. J. and Hahn, F. H.: 1971, *General Competitive Analysis*, Holden-Day, San Francisco, CA.
- Arrow, K. J. and Scitovsky, T.: 1969, *Readings in Welfare Economics*, George Allen and Unwin Ltd., London.
- Baldwin, R. E. and Forslid, R.: 1997, The core-periphery model and endogenous growth. CEPR working paper # 1749, CEPR, London, UK.
- Barro, R. J. and Sala-i-Martin, X.: 1995, *Economic Growth*, McGraw-Hill, New York.
- Buchanan, J. M. and Yoon, Y. J. (eds): 1994, *The Return to Increasing Returns*, The University of Michigan Press, Ann Arbor.
- Cass, D.: 1965, Optimum growth in an aggregate model of capital accumulation, *Review of Economic Studies* **32**, 233–40.
- Chamberlin, E. H.: 1933, *The Theory of Monopolistic Competition*, 1st edn, Harvard University Press, Cambridge, MA.
- Chamberlin, E. H.: 1956, *The Theory of Monopolistic Competition*, 7th edn, Harvard University Press, Cambridge, MA.
- Chipman, J. S.: 1970, External economies of scale and competitive equilibrium, *Quarterly Journal of Economics* **84**(3), 347–85.
- Davis, D. R. and Weinstein, D. E.: 1998a, Does economic geography matter for international specialization? mimeo, School of Business Administration, University of Michigan, Ann Arbor, MI. A slightly different version is available as NBER working paper # 5706.
- Davis, D. R. and Weinstein, D. E.: 1998b, Economic geography and regional production structure: An empirical investigation. Federal Reserve Bank of New York Staff Report # 40.
- Dean, R. D., Leahy, W. H. and McKee, D. L.: 1970, *Spatial Economic Theory*, The Free Press, New York, NY.
- Dixit, A. K. and Stiglitz, J. E.: 1977, Monopolistic competition and optimum product diversity, *American Economic Review* **67**(3), 297–308.
- Domar, E. D.: 1946, Capital expansion, rate of growth, and employment, *Econometrica* **14**, 137–47. Reprinted in Stiglitz and Uzawa (1969).
- Englmann, F. C. and Walz, U.: 1995, Industrial centers and regional growth in the presence of local inputs, *Journal of Regional Science* **35**(1), 3–27.
- Ethier, W. J.: 1982, National and international returns to scale in the modern theory of international trade, *American Economic Review* **72**(3), 389–405.
- Fujita, M.: 1986, Urban land use theory, in J. Lesourne and H. Sonnenschein (eds), *Location Theory*, Harwood Academic Publishers, Chur, CH, pp. 73–149.
- Fujita, M.: 1988, A monopolistic competition model of spatial agglomeration, *Regional Science and Urban Economics* **18**(1), 87–124.

- Fujita, M. and Thisse, J.-F.: 1996, Economics of agglomeration, *Journal of the Japanese and International Economies* **10**, 339–78.
- Gabszewicz, J. and Thisse, J.-F.: 1986, Spatial competition and the location of firms, in J. Lesourne and H. Sonnenschein (eds), *Location Theory*, Harwood Academic Publishers, Chur, CH., pp. 1–71.
- Greenhut, M. L.: 1952, Integrating the leading theories of plant location, *The Southern Economic Journal* **18**, 526–38. Reprinted in Dean, Leahy and McKee (1970).
- Greenhut, M. L.: 1956, *Plant Location in Theory and Practice*, University of North Carolina Press, Chapel Hill, NC.
- Grossman, G. M. and Helpman, E.: 1991, *Innovation and Growth in the Global Economy*, MIT Press, Cambridge, MA.
- Harrod, R.: 1939, An essay in dynamic theory, *The Economic Journal* **49**, 14–33. Reprinted in Stiglitz and Uzawa (1969).
- Heijdra, B. J.: 1997, Toch een revolutie?, *Tijdschrift voor Politieke Economie* **20**(1), 7–33.
- Helpman, E.: 1984, Increasing returns, imperfect markets, and trade theory, in R. W. Jones and P. B. Kenen (eds), *Handbook of International Economics*, North-Holland, Amsterdam, NL, chapter 7, pp. 325–65.
- Helpman, E. and Krugman, P. R.: 1985, *Market structure and foreign trade*, MIT Press, Cambridge, MA.
- Hicks, J.: 1989, The assumption of constant returns to scale, *Cambridge Journal of Economics* **13**, 9–17.
- Hirschman, A.: 1958, *The Strategy of Economic Development*, Yale University Press, New Haven, CT.
- Hotelling, H.: 1929, Stability in competition, *The Economic Journal* **39**, 41–57. Reprinted in Dean et al. (1970).
- Jacobs, J.: 1969, *The Economy of Cities*, Penguin Books, Harmondsworth, UK.
- Jones, C. I.: 1995a, R & D-based models of economic growth, *Journal of Political Economy* **103**(4), 759–84.
- Jones, C. I.: 1995b, Time series tests of endogenous growth models, *Quarterly Journal of Economics* **110**(2), 495–525.
- Judd, K. L.: 1985, On the performance of patents, *Econometrica* **53**(3), 567–585.
- Junius, K.: 1996, Economic development and industrial concentration: An inverted U-curve. Kiel Working Paper # 770, Kiel Institute of World Economics, Kiel, Germany.
- Kaldor, N.: 1970, The case for regional policies, *Scottish Journal of Political Economy* **17**, 337–47.
- Knaap, T.: 1997, Long term convergence in a long panel data set. Paper presented

- at the SOM workshop on Growth, Convergence and International Trade, Groningen University.
- Koopmans, T. C.: 1965, On the concept of optimal economic growth, (*Study Week on the) Econometric Approach to Development Planning*, North-Holland Publishing Co., Amsterdam and Rand McNally, Chicago, chapter 4, pp. 225–87.
- Krugman, P. R.: 1979, Increasing returns, monopolistic competition, and international trade, *Journal of International Economics* **9**(4), 469–479. Reprinted as Chapter 1 in Krugman (1990).
- Krugman, P. R.: 1980, Scale economies, product differentiation, and the pattern of trade, *American Economic Review* **70**, 950–959. Reprinted as Chapter 2 in Krugman (1990).
- Krugman, P. R.: 1990, *Rethinking International Trade*, The MIT Press, Cambridge, MA.
- Krugman, P. R.: 1991a, *Geography and Trade*, Leuven University Press, Leuven, Belgium and MIT Press, Cambridge, MA.
- Krugman, P. R.: 1991b, Increasing returns and economic geography, *Journal of Political Economy* **99**(3), 483–99.
- Krugman, P. R.: 1995, *Development, Geography, and Economic Theory*, The Ohlin Lectures, MIT Press, Cambridge, MA.
- Krugman, P. R.: 1998, Space: The final frontier, *Journal of Economic Perspectives* **12**(2), 161–74.
- Krugman, P. R. and Venables, A. J.: 1995, Globalization and the inequality of nations, *Quarterly Journal of Economics* **110**(4), 857–80.
- Lancaster, K.: 1979, *Variety, Equity and Efficiency*, Columbia University Press, New York, NY.
- Lösch, A.: 1967, *The Economics of Location*, 2nd, rev. edn, John Wiley & Sons Science Editions, New York, NY. Translated from German by W. H. Woglom for Yale University Press.
- Lucas, R. E.: 1988, On the mechanics of economic development, *Journal of Monetary Economics* **22**, 3–42.
- Maddison, A.: 1982, *Phases of Capitalist Development*, Oxford University Press, Oxford (UK).
- Mankiw, N., Romer, D. and Weil, D. N.: 1992, A contribution to the empirics of economic growth, *Quarterly Journal of Economics* **107**, 407–37.
- Marshall, A.: 1920, *Principles of Economics*, Macmillan, London.
- Martin, P. and Ottaviano, G. I. P.: 1996a, Growing locations: Industry location in a model of endogenous growth. CEPR discussion paper # 1523, CEPR, London.

- Martin, P. and Ottaviano, G. I. P.: 1996b, Growth and agglomeration. CEPR working paper # 1529, CEPR, London, UK.
- Matsuyama, K.: 1993, Modelling complementarity in monopolistic competition, *Bank of Japan Monetary and Economic Studies* **11**(1), 87–109.
- Matsuyama, K.: 1995, Complementarities and cumulative processes in models of monopolistic competition, *Journal of Economic Literature* **33**, 701–29.
- Mulligan, C. B. and Sala-i-Martin, X.: 1993, Transitional dynamics in two-sector models of endogenous growth, *Quarterly Journal of Economics* **108**(3), 739–73.
- Myrdal, G.: 1957, *Economic Theory and Underdeveloped Regions*, Duckworth, London.
- Nonneman, W. and Vanhoudt, P.: 1996, A further augmentation of the Solow model and the empirics of economic growth for OECD countries., *Quarterly Journal of Economics* **111**(3), 943–53.
- Ottaviano, G. I. P. and Puga, D.: 1997, Agglomeration in the global economy: A survey of the 'new economic geography'. Discussion Paper # 356, Centre for Economic Performance, London, UK.
- Pack, H.: 1994, Endogenous growth theory: Intellectual appeal and empirical shortcomings, *Journal of Economic Perspectives* **8**(1), 55–72.
- Rebelo, S.: 1991, Long-run policy analysis and long-run growth, *Journal of Political Economy* **99**(3), 500–21.
- Rensman, M.: 1996, Economic growth and technological change in the long run. SOM Research Report 96C10, Groningen University, Groningen, The Netherlands.
- Rivera-Batiz, F. L.: 1988, Increasing returns, monopolistic competition, and agglomeration economies in consumption and production, *Regional Science and Urban Economics* **18**(1), 125–53.
- Rivera-Batiz, L. A. and Romer, P. M.: 1991, Economic integration and endogenous growth, *Quarterly Journal of Economics* **106**(2), 531–55.
- Robinson, J.: 1933, *The economics of imperfect competition*, 1st edn, MacMillan, London, UK.
- Romer, D.: 1993, The new keynesian synthesis, *Journal of Economic Perspectives* **7**(1), 5–22.
- Romer, P. M.: 1986, Increasing returns and long-run growth, *Journal of Political Economy* **94**(5), 1002–37.
- Romer, P. M.: 1987, Growth based on increasing returns due to specialization, *AEA Papers and Proceedings* **77**(2), 56–62.
- Romer, P. M.: 1990, Endogenous technological change, *Journal of Political Economy* **98**(5, pt. 2), S71–102.

- Romer, P. M.: 1991, Increasing returns and new developments in the theory of growth, in W. A. Barnett, B. Cornet, C. D'Aspremont, J. Gabszewicz and A. Mas-Colell (eds), *Equilibrium Theory and Applications: Proceedings of the Sixth International Symposium in Economic Theory and Econometrics*, Cambridge University Press, pp. 83–110.
- Romer, P. M.: 1994, The origins of endogenous growth, *Journal of Economic Perspectives* **8**(1), 3–22.
- Schumpeter, J. A.: 1942, *Capitalism, Socialism and Democracy*, Harper & bros., New York, NY.
- Scitovsky, T.: 1954, Two concepts of external economies, *Journal of Political Economy* **62**, 143–51.
- Smith, A.: 1776, *An Inquiry into the Nature and Causes of the Wealth of Nations*, Clarendon Press, Oxford.
- Solow, R. M.: 1956, A contribution to the theory of economic growth, *Quarterly Journal of Economics* **70**, 65–94. Reprinted in Stiglitz and Uzawa (1969).
- Solow, R. M.: 1994, Perspectives on growth theory, *Journal of Economic Perspectives* **8**(1), 45–54.
- Spence, M.: 1976, Product selection, fixed costs, and monopolistic competition, *Review of Economic Studies* **43**, 217–35.
- Starrett, D.: 1978, Market allocation of location choice in a model with free mobility, *Journal of Economic Theory* **17**, 21–37.
- Stigler, G. J.: 1951, The division of labor is limited by the extent of the market, *Journal of Political Economy* **59**(3), 185–93. Reprinted in Buchanan and Yoon (1994).
- Stiglitz, J. E. and Uzawa, H.: 1969, *Readings in the Modern Theory of Economic Growth*, The MIT Press, Cambridge, MA and London, UK.
- Summers, R. and Heston, A.: 1988, A new set of international comparisons of real product and price levels estimates for 130 countries, *Review of Income and Wealth* **34**, 1–26.
- Swan, T. W.: 1956, Economic growth and capital accumulation, *The Economic Record* **32**(63), 334–61. Reprinted in Stiglitz and Uzawa (1969).
- Takayama, A.: 1985, *Mathematical Economics*, 2nd edn, Cambridge University Press, New York, NY.
- Venables, A. J.: 1996, Equilibrium locations of vertically linked industries, *International Economic Review* **37**(2), 341–359.
- Von Thünen, J. H.: 1842, *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*, Gustav Fischer Verlag, Jena.
- Weber, A.: 1909, *Über den Standort der Industrien*, Mohr Verlag, Tübingen.

- Weitzman, M. L.: 1994, Monopolistic competition with endogenous specialization, *Review of Economic Studies* **61**, 45–56.
- Young, A. A.: 1928, Increasing returns and economic progress, *The Economic Journal* **38**, 527–42. Reprinted in Arrow and Scitovsky (1969).